

University of Groningen

## The Role of Resource Misallocation in Cross-country Differences in Manufacturing Productivity

Inklaar, Robert; Lashitew, Addisu A.; Timmer, Marcel P.

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

### *Document Version*

Publisher's PDF, also known as Version of record

### *Publication date:*

2014

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

Inklaar, R., Lashitew, A. A., & Timmer, M. P. (2014). *The Role of Resource Misallocation in Cross-country Differences in Manufacturing Productivity*. (GGDC Working Papers; Vol. GD-143). GGDC.

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

## **The Role of Resource Misallocation in Cross-country Differences in Manufacturing Productivity**

Robert Inklaar, Addisu A. Lashitew and Marcel P. Timmer

January 2014

university of  
 groningen

groningen growth and  
 development centre

# **The Role of Resource Misallocation in Cross-country Differences in Manufacturing Productivity**

Robert Inklaar, Addisu A. Lashitew and Marcel P. Timmer

Groningen Growth and Development Centre, University of Groningen

January 24, 2014

## **ABSTRACT**

When capital and labor are not allocated to the more productive firms, aggregate total factor productivity (TFP) suffers. Can this explain observed productivity differences across countries? We estimate manufacturing TFP levels for 52 developing countries and decompose it into a part due to misallocation and a part due to (residual) technology differences. The results show that removing misallocation would increase TFP by an average of 60 percent, but productivity gaps relative to the US remain large. The degree of misallocation is uncorrelated with observed productivity.

Lashitew would like to thank the Netherlands Organization for Scientific Research (NWO) for financial support. Address correspondence to Robert Inklaar, Groningen Growth and Development Centre, Faculty of Economics and Business, University of Groningen, P.O. Box 800, 9700 AV Groningen, the Netherlands; e-mail: R.C.Inklaar@rug.nl.

## I. INTRODUCTION

Total factor productivity (TFP) differences are of great importance in accounting for income differences across countries (Caselli, 2005; Hsieh and Klenow 2010). Under the assumption that TFP measures production technology, this would point to the importance of factors such as slow technology adoption in less developed countries (Parente and Prescott 1994; Comin and Hobijn, 2010). However, measured TFP may reflect not only technology but also misallocation of resources caused by distortions in output and factor markets.<sup>1</sup> Since improving the efficiency of resource allocation across firms is likely a very different challenge from improving the technology that firms use, it is important to disentangle these two aspects of measured TFP.

The contribution of this paper is to determine the importance of resource misallocation for cross-country differences in manufacturing productivity. Hsieh and Klenow (2009) have demonstrated that eliminating resource misallocation across manufacturing plants would lead to larger productivity gains in China and India than in the US. In this paper, we investigate the importance of resource misallocation for a much broader sample of countries. We use the World Bank Enterprise Survey (WBES), a standardized survey that contains plant-level financial data for a wide range of developing and emerging economies. Following the Hsieh and Klenow (2009) methodology, we conduct a liberalization experiment to quantify the productivity gains from reducing resource misallocation around the year 2005.<sup>2</sup> This is done using data for 52 countries that span much of the development spectrum, from a GDP per capita level of 0.52 percent of the US level (Democratic Republic of Congo) to 52 percent of the US level (Slovenia).<sup>3</sup> We find that most countries would benefit considerably from reducing the degree of resource misallocation to the level seen in the US, with an average increase in manufacturing TFP of 62 percent.

To put these findings into perspective, we estimate relative manufacturing productivity levels, building on and extending the approach of Herrendorf and Valentinyi (2012). Relative TFP is computed as relative value added per worker divided by relative factor inputs (physical and human capital) per worker. To measure relative value added per worker we estimate relative output prices, using not just prices of consumption and investment goods but also of exports and imports.<sup>4</sup> Relative factor inputs are computed using data on relative wages and rental prices. We find that even if all resource misallocation were eliminated, productivity differences would remain substantial. The average observed productivity level in our set of 52 countries is 23 percent of the US

---

<sup>1</sup> See e.g. Basu and Fernald (2002), Jones (2011, 2013) and Bartelsman, Haltiwanger and Scarpetta (2013).

<sup>2</sup> The WBES surveys have been held in years varying between 2002 and 2010, so 2005 is a central year in this range.

<sup>3</sup> According to the Penn World Table (PWT), version 8.0, for 2005 (Feenstra, Inklaar and Timmer, 2013). Only 37 out of 167 countries have even higher GDP per capita levels.

<sup>4</sup> This follows an approach similar to Inklaar and Timmer (2012) and is consistent with the most recent version of the Penn World Table (Feenstra, et al., 2013).

level and this rises to 37 percent after eliminating misallocation. While this represents a substantial improvement, resource misallocation is not important enough to explain low productivity levels. More importantly, the countries that would gain most from eliminating resource misallocation are not necessarily the ones with the lowest productivity levels.

We establish the robustness of these results by considering various measurement alternatives for misallocation and for manufacturing productivity. We vary assumptions about factor elasticities and the elasticity of substitution, both of which are important in the Hsieh and Klenow (2009) framework. We also consider different data sources and assumptions for manufacturing productivity measurement. Our main findings are robust to these measurement alternatives.

A few remarks are useful in putting these results in a broader context. In the terminology of Restuccia and Rogerson (2013) we follow an indirect approach, whereby the full gap between marginal costs and marginal products of capital and labor is labeled as resource misallocation. This indirect approach contrasts with direct approaches, which analyze the role of a specific friction – such as financial frictions (Buera, Kaboski and Shin, 2011) – on resource allocation and hence aggregate productivity. Compared to such studies, our approach is broader but also less closely tied to specific frictions. Furthermore, by attributing the full gap between marginal costs and marginal products to misallocation, we may be overstating the importance of misallocation: adjustment costs, experimentation by firms with new technologies and measurement error are all included as part of misallocation.

Our focus on within-industry, between-firm variation means that we ignore between-sector misallocation of the sort emphasized by Vollrath (2009) and Fernald and Nieman (2011). Both find that some sectors (agriculture, subsidized manufacturing) may employ an inefficiently large part of the labor force. Our results for manufacturing may also not be representative for the rest of the economy. For instance, Adamopoulos and Restuccia (2014) show how distortions to farm size are systematically able to account for part of the cross-country productivity differences in agriculture. Finally, in the Hsieh and Klenow (2009) approach, the underlying production technology of firms is considered exogenous, but from the broader literature on productivity (e.g. Syverson, 2011) we know that firms can and do engage in technology-enhancing investments, such as spending on research and development. We also know that, for instance, financial frictions can lead to sub-optimal investment in such long-run projects (Aghion, Angeletos, Bannerjee and Manova, 2010), thus leading to a link between factor misallocation and firm technology. Seen in this light, our results serve mostly to indicate how important one specific type of misallocation is for cross-country productivity and (ultimately) income differences.

In the remainder of this paper, we first outline the theoretical framework of the Hsieh and Klenow (2009) model in Section II, before turning to a discussion of how we

measure manufacturing productivity levels and the efficiency of resource allocation in Section III. We present the main results in Section IV, sensitivity analysis in Section V and we provide some conclusion in Section VI.

## II. THEORETICAL FRAMEWORK

In order to identify the contribution of misallocation to cross-country differences in total factor productivity (TFP), we need measures of both TFP and misallocation. Therefore, we first conduct a development accounting analysis to evaluate the contribution of TFP to labor productivity differences across countries. In the second stage, we identify the contribution of misallocation to cross-country productivity differences using measures of misallocation based on firm-level data using the model and methodology proposed by Hsieh and Klenow (2009).

When there are perfect factor and product markets, aggregate productivity reflects only technological differences across countries. But in the presence of distortions that drive a wedge between the marginal product and marginal cost of productions factors, aggregate productivity will also reflect resource misallocation (Basu and Fernald, 2002; Fernald and Neiman, 2011). Hsieh and Klenow (2009) argue that misallocation of resources between firms, within industries can be important in explaining TFP differences across countries. Applying their model to firm-level data we would be able to write actual TFP ( $A$ ) as a ratio of the level of (hypothetical) efficient TFP ( $A^*$ ) and the efficiency of resource allocation ( $RA$ ). The ratio of TFP in country  $c$  relative to country  $k$  ( $A_{ck}$ ) can then be written as follows:

$$(1) \quad A_{ck} = A_{ck}^* \times RA_{ck}$$

When the efficiency of resource allocation is equal between the two countries, aggregate TFP differences are determined only by what we label as technology differences.<sup>5</sup> Measuring the extent to which misallocation reduces aggregate TFP involves: i) calculating the actual level of TFP from plant level data; and ii) calculating the hypothetical TFP that would be achieved if there were no misallocation - i.e. if marginal products are equalized within industries. To illustrate this, we provide a brief sketch of the Hsieh and Klenow (2009) model below.

At the highest level of aggregation, final output  $Y_c$ , is produced by combining output from manufacturing industries  $Y_{c,s}$  using Cobb-Douglas production technology:

$$(2) \quad Y_c = \prod_{s=1}^S (Y_{c,s})^{\theta_{c,s}},$$

where  $\theta_{c,s}$  is the value added share of sector  $s$  in country  $c$ , and  $S$  is the total number of manufacturing industries. Industry output  $Y_s$  (omitting country subscripts for simplicity) is a CES aggregate of  $M_s$  differentiated products:

---

<sup>5</sup> Note that the efficiency of resource allocation in equation (1) is the inverse of the TFP gains metric presented by Hsieh and Klenow's (2009).

$$(3) \quad Y_s = \sum_{i=1}^{M_s} \left( Y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$

where  $Y_{si}$  is a differentiated product by firm  $i$  in industry  $s$ , and  $\sigma$  is the elasticity of substitution. Each differentiated product is produced by firms with heterogeneous productivity ( $A$ ) using labor ( $L$ ) and capital ( $K$ ) with Cobb-Douglas technology:

$$(4) \quad Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}.$$

The main feature of the model is that firms are not only heterogeneous with respect to their productivity, as in Melitz (2003), but they also face idiosyncratic distortions to their input and output prices. Two types of distortions are introduced: *output distortions* that affect the quantity of production while leaving the input mix unaffected, and *capital distortions* that affect the use of capital relative to labor. The output distortion is modeled as a tax on production – independent of factor use – because it distorts the marginal products of capital and labor in equal proportions. Capital distortion, on the other hand, is a form of tax on capital and thus affects the input mix decision. Note that both distortions are exogenous and are implied from the data as discussed below.

In this framework, profits depend not only on prices and quantities, but also on distortions:

$$(5) \quad \pi_{si} = P_{si}(1 - \tau_{Ysi})Y_{si} - wL_{si} - (1 + \tau_{Ksi})rK_{si},$$

where  $w$  is the wage rate,  $r$  is the rental price of capital,  $\tau_{ysi}$  is output distortion and  $\tau_{ksi}$  is capital distortion. Profit maximization leads to the standard condition that the firm's output price is a fixed markup over its marginal cost:

$$(6) \quad P_{si} = [\sigma/(\sigma - 1)][w/(1 - \alpha_s)]^{1-\alpha_s} [r/\alpha_s]^{\alpha_s} [(1 + \tau_{Ksi})^{\alpha_s}/(1 - \tau_{Ysi})A_{si}],$$

where the term  $(\sigma/(\sigma - 1))$  is the markup of prices over marginal costs. In addition to factor prices, both output and capital distortions appear in the price equation with a positive effect. The marginal revenue product of labor (MRPL) and the marginal revenue product of capital (MRPK) are given by the respective partial derivatives of the revenue function multiplied by the inverse of the markup to correct for rents:

$$(7) \quad MRPL_{si} \equiv [1 - \alpha_s][(\sigma - 1)/\sigma][P_{si}Y_{si}/L_{si}] = w/(1 - \tau_{Ysi}).$$

$$(8) \quad MRPK_{si} \equiv \alpha_s[(\sigma - 1)/\sigma][P_{si}Y_{si}/K_{si}] = [r(1 + \tau_{Ksi})/(1 - \tau_{Ysi})].$$

Equations (7) and (8) show that the marginal revenue products of labor and capital are determined not only by the wage rate and the rental price of capital but also by distortions. Capital distortions raise only the marginal revenue of capital whereas output distortions raise both the marginal revenue product of labor and capital. To link the two measures of distortion with aggregate productivity, it is important to note the

distinction between revenue TFP,  $TFPR$ , and quantity TFP,  $TFPQ$ :  $TFPQ$  is a measure of total factor productivity after accounting for firm-level price differences, whereas  $TFPR$  is a measure of productivity that is not separated from price (i.e.  $TFPR = TFPQ \times P$ ).

$$(9) \quad TFPQ_{si} \equiv A_{si} = Y_{si}/K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}.$$

$$(10) \quad TFPR_{si} \equiv P_{si} A_{si} = P_{si} Y_{si}/K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}.$$

By using price equation (6),  $TFPR$  can be expressed as a function of distortions and factor prices. Since all distortions are reflected in factor marginal products,  $TFPR$  can also be alternatively expressed as a function of the marginal revenue products of capital and labor:

$$(11) \quad TFPR_{si} = \left( \frac{\sigma}{\sigma-1} \right) \left[ \frac{w}{1-\alpha_s} \right]^{1-\alpha_s} \left[ \frac{r}{\alpha_s} \right]^{\alpha_s} \left[ \frac{(1+\tau_{Ksi})^{\alpha_s}}{1-\tau_{Ysi}} \right] = \left[ \frac{MRPL_{si}}{1-\alpha_s} \right]^{1-\alpha_s} \left[ \frac{MRPK_{si}}{\alpha_s} \right]^{\alpha_s}$$

Equation (11) shows that all differences in  $TFPR$  within an industry are caused by output and capital distortions. Note that no physical productivity ( $TFPQ$ ) term features in the equation, and thus  $TFPR$  has no relationship with physical/quantity productivity. Although firms with high physical productivity ( $TFPQ$ ) have high revenue productivity by definition (equation (10)), they also charge lower prices since they are cost efficient (equation (6)). This relationship allows us to use  $TFPR$  to capture the effects of both types of distortions. Similarly, industry-level revenue productivity  $TFPR_s$  can be shown to be a function of distortions:

$$(12) \quad TFPR_s = \left( \frac{\sigma}{\sigma-1} \right) \left[ \frac{w}{1-\alpha_s} \sum_i (1-\tau_{Ysi}) \eta_{si} \right]^{1-\alpha_s} \left[ \frac{r}{\alpha_s} \sum_i \left( \frac{1-\tau_{Ysi}}{1+\tau_{Ksi}} \right) \eta_{si} \right]^{\alpha_s}$$

where the weighting term  $\eta_{si}$  is the output share of firm  $i$  in industry  $s$ . Industry productivity is given by the following equation:

$$(13) \quad A_s = \left[ \sum_{i=1}^{M_s} (A_{si} TFPR_s / TFPR_{si})^{\sigma-1} \right]^{\frac{1}{\sigma-1}},$$

where  $A_{si}$  is physical productivity ( $TFPQ$ ), and  $TFPR_s$  is industry-level revenue productivity. Without resource misallocation, so if all firms would face zero output and capital distortion, industry-level TFP (equation (13)) would be fully determined by firm productivity:

$$(14) \quad A_s^* = \left[ \sum_i A_{si}^{\sigma-1} \right]^{\frac{1}{\sigma-1}}$$

After aggregating across industries using industry value added, we get a country-level measure of the efficiency of resource allocation in manufacturing (cf. equation (1)):



$$(15) \quad RA_c = A_c / A_c^*$$

### III. DATA AND MEASUREMENT

#### 3.1 Manufacturing TFP levels

Computing manufacturing TFP levels requires an estimate of manufacturing value added per worker and an estimate of factor inputs per worker. The ratio of the two is then manufacturing TFP. Manufacturing value added in national currency is available from UN National Accounts data, but to make this comparable across countries, we need relative prices of manufacturing output.<sup>6</sup> We assume zero economic profits, so that manufacturing value added equals the payments to labor and capital. To make this comparable across countries, we need the relative prices of labor and capital and the factor elasticities to combine these into an overall factor inputs price.

##### A: Manufacturing output prices

Ideally, relative output price estimates would be based on producer price data, but the lack of dedicated survey data means that a variety of approaches have been followed in the literature. When focused only on manufacturing, some have opted to use exchange rates to compare output from different countries, assuming a relative price of one (e.g. Rodrik, 2013). An argument in favor of this approach is that many manufactured products are traded and thus more exposed to the pressures of the Law of One Price (LOP). But this argument is not fully convincing given the systematic deviations from LOP even for products that are internationally traded (Feenstra and Romalis, 2012; Burstein and Gopinath, 2013) and the very limited trade in some manufactured products, such as ready-mixed concrete (Syverson, 2008).

The main alternative approach is to use relative prices collected as part of the International Comparison Program (ICP). These price form the basis of the GDP PPPs disseminated by the World Bank (2008) and are expenditure prices of consumption and investment goods and services. Relative output prices for manufacturing are then estimated by selecting and combining the prices of goods that are made by manufacturing industries, as in Sørensen and Schjerning (2008), Van Biesebroeck (2009) and Herrendorf and Valentinyi (2012). Given its broad application, it can be seen as the standard approach.

Yet this standard approach has drawbacks as well. Most importantly, the prices of goods consumed or invested domestically do not take into account the prices of exported products while the prices of imported goods are included. This problem is compounded by relying on the value of consumption and investment expenditure to aggregate more detailed prices, rather than using the value of output. As detailed in the appendix, we remedy both problems here. We combine ICP data on consumption and investment prices and expenditure (used in the standard approach) with data on industry output,

---

<sup>6</sup> In the appendix, we also detail how the estimation of the number of manufacturing workers.

exports and imports and relative prices of exports and imports from Feenstra and Romalis (2012).

### B: Input prices

To compute manufacturing productivity levels, we need prices of inputs in addition to the price of manufacturing output. Ideally, an overall input price index should be compiled using prices of capital, labor and intermediate inputs. However, an intermediate inputs price index requires the type of detailed input-output data that is mostly missing for the set of countries we analyze here. So instead, we assume the price for manufacturing output equals the price for manufacturing value added. The results of Inklaar and Timmer (2013) for 42 countries provide some support for this assumption. They do estimate separate output and intermediate input prices and for manufacturing as a whole, the correlation between the output and value added prices is very high and their variance is similar.<sup>7</sup>

That leaves estimating the relative price of labor and of capital. Estimating relative wages is challenging as the aim is to measure the wage of the same type of worker in different countries (Ashenfelter, 2012). Differences in educational qualifications or differences in occupational composition and characteristics can all stand in the way of identifying the ‘same type’. A related issue, which is particularly relevant in the current context, is that the ‘same type’ of worker may earn a different wage in different sectors.<sup>8</sup> Finally, we want to compare the full cost of employing a worker, labor compensation, which includes both the wage they earn as well as any contributions to social security or other benefits. Given that our aim is to compare productivity across a group of countries that includes a number without an extensive statistical infrastructure, we inevitably have to compromise between these goals.

Our wage measure for the majority of countries is based on the same principle as Herrendorf and Valentinyi (2012), namely the country-average wage level adjusted for differences in schooling.<sup>9</sup> The data source for this measure is the Penn World Table, version 8.0 (see Feenstra, Inklaar and Timmer, 2013) and this wage measures covers all of labor compensation.<sup>10</sup> For a few countries, we use economy-wide wages, not adjusted for differences in schooling. For the remainder of countries, we compute the median manufacturing wage from the World Bank Enterprise Survey (WBES), based on labor compensation and the number of workers of each manufacturing firm. Below, we also show results relying solely on WBES wages. From all these sources, we use data for 2005 or the nearest available year in case of the WBES) To put the countries on a

---

<sup>7</sup> The correlation of log output and log value added prices is 0.83, the variance of log output prices is 0.048 and the variance of log value added prices is 0.033. In comparison, if we apply the current method of computing output prices to the same set of countries, the variance of log output prices is 0.037.

<sup>8</sup> The variation of wages across firms within the same industry is used for determining the degree of misallocation.

<sup>9</sup> Herrendorf and Valentinyi (2012) assume that the share of each sector in total labor input equals the share in labor compensation, which is equivalent to assuming the same wage across sectors.

<sup>10</sup> Specifically, we multiply exchange-rate converted GDP at current prices by PWT’s labor share in GDP and divide by the number of workers times the human capital index relative to the USA.

comparable basis, we use the trend in overall inflation (of the GDP deflator) to estimate 2005 wage levels for all countries.

The relative price of capital input is computed as the relative rental price. The concept is based on Hall and Jorgenson (1967), as adapted by Jorgenson and Nishimizu (1978) for cross-country comparisons. The relative rental price  $p^k$ , aggregated over  $A$  assets, is computed as:

$$(16) \quad \log(p_j^k) - \overline{\log(p^k)} = \frac{1}{2} \sum_{a=1}^A (c_{aj} + \bar{c}_a) \left[ \log(uc_{aj} p_{aj}^i) - \overline{\log(uc_a p_a^i)} \right],$$

where  $uc_{aj} = i_j + \delta_a - dp_{aj}$  is the user cost of capital with  $i_j$  the nominal interest rate, here taken as the lending rate from the IMF's International Financial Statistics,<sup>11</sup>  $\delta_a$  the asset-specific geometric depreciation rate, and  $dp_{aj}$  the price change of asset  $a$  in country  $j$ . In this expression, a bar over a variable indicates the arithmetic mean across countries. This means that each country is compared to a (hypothetical) average country to ensure that the resulting relative price measure does not depend on the base country that is chosen (see Caves, Christensen and Diewert, 1982a).

The use of the lending rate means we rely on an external rate of return. The alternative would be to choose the rate of return to exhaust the fraction of GDP not paid out as labor compensation. Such an internal rate of return has a number of practical drawbacks (see Inklaar, 2010), but more importantly, we do not have the data on labor compensation for all countries in the analysis. The user cost, relative to the cross-country average, is multiplied by the relative investment price  $p_{aj}^i / \bar{p}_a^i$ . The relative rental price is aggregated across assets using the share of each asset in capital compensation:

$$(17) \quad c_{aj} = kc_{aj} / \sum_a kc_{aj}, \text{ where } kc_{aj} = uc_{aj} p_{aj}^i K_{aj}$$

where  $K_{aj}$  is the capital stock of asset  $a$  in country  $j$ . Capital stocks, asset deflators and depreciation rates are the same as used for the Penn World Table, version 8.0, and these data are described in detail in Inklaar and Timmer (2013). Capital stocks are built up from investment by asset using the perpetual inventory method, based on time series going back as far as 1950. These investment series are partly taken from the OECD National Accounts database and EU KLEMS, and partly estimated based on ICP expenditure data and the commodity flow method.

The relative rental price defined in equation (16) depends on a wide range of data: capital stocks by assets, proper deflators, and interest rates in addition to relative prices

---

<sup>11</sup> If the lending rate is missing, the yield on treasury bonds or bills (also from the International Financial Statistics) is used.

of investment goods and may thus be sensitive to measurement errors in any of these. As we show below, though, the final results are not sensitive to whether we use our preferred rental price measure or a simpler measure of relative investment prices, weighted using investment shares  $w$ :

$$(18) \quad \log(p_j^i) - \overline{\log(p^i)} = \frac{1}{2} \sum_{a=1}^A (w_{aj} + \bar{w}_a) (\log(p_{aj}^i) - \overline{\log(p_a^i)})$$

### C: TFP Calculation

The last piece of information we need for computing relative productivity is elasticity parameters for weighting the prices of labor and capital. We assume that the output elasticities of capital and labor are well-approximated by their US cost share. This reflects the Hsieh and Klenow (2009) model, whereby variations in observed factor shares relative to (assumed) output elasticities reflect misallocation of resources.<sup>12</sup> We use the cost shares as published by the US Bureau of Labor Statistics (BLS) as part of the Major Sector Multifactor Productivity. Those data show that the share of capital income in manufacturing value added is 40.6 percent, with the remainder going to labor. The BLS capital share also covers capital income from land and inventories, so it represents the full contribution of capital to value added.<sup>13</sup> Note that this capital share is higher than the 33 percent of Valentinyi and Herrendorf (2008, Table 1) based on the 1997 US Input-Output table. This is in part due to the increase in the capital share between 1997 and 2005, from 37.4 to 40.6 percent in the BLS data. This is in line with the evidence of an increase in the US capital share of Elsby, Hobijn and Şahin (2013) and fits the broader global upward trend of the capital share, that is analyzed in Karabarbounis and Neiman (2014). Further differences could be due to the focus of Valentinyi and Herrendorf (2008) on income shares in producing manufacturing products rather than on income shares of firms in manufacturing, which is a more natural unit of analysis for our purposes.

### 3.2 Measuring misallocation

The main data source for this analysis is the World Bank's Enterprises Survey (WBES), an ongoing survey that collects firm-level data worldwide. The major advantage of the WBES survey is that data collection is conducted systematically using standardized survey instruments. The dataset thus provides comparable data that is unique in its extensive country coverage. Sampling for the WBES is conducted using stratified sampling procedure to ensure representativeness. First, the number industry groups to be covered across each major sector (services, manufacturing and non-agriculture primary activities) is determined. For manufacturing, industry grouping is based on 2-digit ISIC classification. The number of industry groups to be covered in each country is

---

<sup>12</sup> Besides Hsieh and Klenow (2009), Restuccia and Rogerson (2008) and Fernald and Nieman (2011) also use US cost shares as a (relatively) undistorted measure of output elasticities.

<sup>13</sup> Under the assumption that non-agricultural land represents a constant fraction of 24% of the fixed reproducible capital stock, following the estimate of World Bank (2006), the relative input level is not affected.

determined according to the size of the total economy which is taken as a proxy for the universe of firms.

Once the number of industries is decided, industry groups that contribute relatively more to the total economy in terms of total production or employment are selected. In the second stage, a sampling equation is used to determine a representative sample size per industry group. The sample size is decided with the aim of arriving at a representative sample for the proportion of firms and the average sales in the industry. Finally, further stratification is made based on firm-size and geographical location to select the firms that are covered by the survey.<sup>14</sup>

Data collection started in 2002 and different countries have been covered in subsequent years. Panel data is available for some countries; however, the country coverage of the panel dataset is limited. For the analysis in this paper, we construct a cross-section dataset for coverage of a maximum number of countries. When multiple years of data are available for a country, we use data for the year with the largest number of firm observations.

We started compiling the cross-section data by removing non-manufacturing firms and observations with missing or incomplete data on total production, cost of intermediate inputs, capital stock and labor inputs. Market value of production is not available for most firms, and so the more widely available data of total sales is used. Value added is measured as the difference between sales and the cost of intermediate inputs. Cost of intermediate inputs is calculated by adding up three major cost categories: energy consumption (fuel, electricity and other energy costs), cost of raw materials and overhead and other expenses. To account for differences in hours worked and human capital, we use labor cost rather than employment as a measure of labor inputs. Loss-making firms with negative value added were removed. Then we remove outliers that are likely to be measured with error and can also significantly influence the measures of misallocation. For this purpose, we follow Hsieh and Klenow (2009) and remove the top and bottom percentiles of the two types of distortions as well as of total factor productivity (TFPQ) within each country dataset.

Once the data is cleaned, a number of industries end up with too few valid observations compared to the original sample in which no cleaning is made. The largest loss of data is caused by lack of capital stock data. To make sure that the final sample is not too different from the original sample, which is designed to be representative, we exclude industries if they have fewer than five observations, or if the number of usable observation is less than half the number of original observations.

The exclusion of certain industries in this way leaves many countries with too few observations. We exclude all countries which have fewer than 40 observations, and

---

<sup>14</sup> A full description of the sampling procedure can be found at [www.enterprisesurveys.org/methodology](http://www.enterprisesurveys.org/methodology).

whose sample in terms of coverage relative to the original sample is less than 40%. While the decision for the cut-off point is rather arbitrary, it ensures the exclusion of countries in which the final dataset is not likely to be representative. This leads to a final dataset of 52 countries with a total of 20,378 plants.

Table A2 in the appendix lists the countries covered in our dataset and the number of firm observations per country. The average sample size across countries is close to 400, although there is also large difference in sample size across countries. Whereas large countries such as India, Brazil and China have well above a thousand observations, smaller ones such as Estonia and Swaziland have only around 40 observations. The dataset covers 52 mostly low- and middle-income countries with a median per capita GDP \$3164 in 2005 (in PPP-converted US dollars from PWT 8.0).<sup>15</sup> The country with the highest income is Slovenia (\$21967) and the one with the lowest income is the Democratic Republic of Congo (\$221).

### 3.3 Measures of misallocation

A number of parameters are required to calculate the efficiency of resource allocation as given by equation (15). First, we need to specify values for the wage rate and the rental price of capital in order to measure the marginal products of labor and capital. For every country in the dataset, we set the wage rate to the average value of the observed wage rate among firms within the country. For all countries, we set the rental price of capital  $r$  to 0.10, assuming a real interest rate of 5% and a depreciation rate of 5% as in Hsieh and Klenow (2009). Incorrectly measuring the wage rate and the rental price of capital does not affect our measures of TFP gap. This is because the error will be reflected in the marginal products of labor and capital of all firms, thus affecting the distortion of all firms in equal amount.

Second, we need to assign a value for the elasticity of substitution ( $\sigma$ ) among products. As can be seen from the efficient TFP in equations (13) and (14), the elasticity of substitution affects the level of actual and efficient industry TFP and hence the TFP gap. Again, we follow Hsieh and Klenow (2009) and choose an elasticity parameter of 3,<sup>16</sup> though we also experiment with the higher value of 5 in the sensitivity analysis.

Finally, benchmark values of the output elasticity of capital and labor are required in order to measure distortions. It is necessary to apply similar industry-specific elasticity parameters for all countries in our dataset in order to get comparable measures of distortions and TFP gap. These benchmark elasticity parameters should come from data that are not distorted and thus reflect the true characteristics of each industry's technology. Again following the precedence of Hsieh and Klenow (2009), we use elasticity parameters from the relatively less distorted US economy as benchmark

---

<sup>15</sup> This compares with a median level of GDP per capita of \$6573 across all 167 countries in PWT.

<sup>16</sup> Based on the median elasticity of substitution estimated by Broda and Weinstein (2006) for the most recent period of time of 3.1.

values. Table A3 in the appendix provides the elasticity parameters used for our analysis at 2-digit level industrial classification.

Once these parameters are determined, output and capital distortions can be computed based on the model. Using the definition of MRPL from equation (7), the output distortion of a firm is measured as the gap between its labor share (multiplied by the markup to adjust for rents) and the labor share of a representative US firm in the same industry:

$$(19) \quad \tau_{Ysi} = 1 - \left( \frac{\sigma}{\sigma - 1} \right) \left( \frac{wL_{si} / P_{si} Y_{si}}{1 - \alpha_s} \right)$$

If a firm faces a high MRPL, this will show up as a low labor share in value added for a given wage rate. This lowers the ratio of the firm's labor share to the labor share of the representative US firm, reflecting a high output distortion. Using the definitions of MRPL and MRPK given by equations (7) and (8), the capital distortion is computed from the gap between the firm's capital-labor ratio and the capital-labor ratio of the US industry-representative firm:

$$(20) \quad \tau_{Ksi} = \left( \frac{\alpha_s}{1 - \alpha_s} \right) \left( \frac{wL_{si}}{rK_{si}} \right) - 1$$

The implication here is that if a firm has a lower capital-labor ratio compared to the US industry benchmark, it is facing higher capital distortions. Based on the Cobb-Douglas technology assumed in equation (3), firm-level productivity is measured as:

$$(21) \quad A_{si} = \frac{(P_{si} Y_{si})^{\frac{\sigma}{\sigma-1}}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}}$$

This equation enables us to measure physical productivity (*TFPQ*) by deriving quantities from revenues using a demand function that establishes the relationship between quantity and prices. The exponent in the numerator of equation (21) is the derivation of the elasticity parameter that is used to convert revenues to quantities. Once productivity and distortions are calculated, we are able to measure the efficiency of resource allocation as the ratio of observed to the (undistorted) efficient TFP.

#### IV. ANALYSIS

This provides us with all the necessary inputs to determine the role of resource allocation in manufacturing productivity differences. To first provide some perspective on the role of manufacturing productivity differences, Figure 1 plots our measure of manufacturing TFP against manufacturing value added per worker for the 52 countries

in our analysis.<sup>17</sup> The graph shows that TFP is strongly correlated with manufacturing labor productivity, but also that the variation in observed labor productivity is much larger than the variation in TFP levels, as indicated by the scale of the axes. The variation in TFP levels is approximately one-third of the variation in value added per worker. In other words, part of the variation in value added per worker can be accounted for through the variation in factor inputs per worker. The relative variation of TFP compared to the variation in labor productivity is broadly comparable to results for the aggregate economy for the same year, 2005 (see Feenstra et al. 2013). This result is in line with the finding of Herrendorf and Valentinyi (2012) that the variation in manufacturing TFP is of a similar magnitude as the variation in economy-wide TFP.

**Figure 1 Manufacturing TFP and labor productivity relative to the US.**



We now consider the efficiency of resource allocation and how this affects observed TFP differences. Figure 2 presents the first result by comparing observed TFP with efficient TFP, i.e. the TFP level that would be attained if all distortions are removed. This figure illustrates that TFP gains from removing distortions are substantial, with the average TFP level relative to the US increasing from 23 to 37 percent. However, for all but a few countries, the TFP differences would remain large after removing these distortions. To illustrate, in the observed TFP data only four countries have a TFP level that is  $\frac{1}{2}$  of the US level or higher, while after removing distortions, 12 countries pass this level.<sup>18</sup> So while removing distortions would be beneficial for productivity, this would not – by itself – be enough to eliminate productivity differences.

<sup>17</sup> The TFP results by country are shown in Table A2 in the Appendix.

<sup>18</sup> Indeed, Croatia (HRV) would even reach a TFP level that is 37 percent higher than in the US. That specific finding, though, is sensitive to the precise wage and rental price data used.



**Figure 2 Observed and efficient TFP (USA=1)**

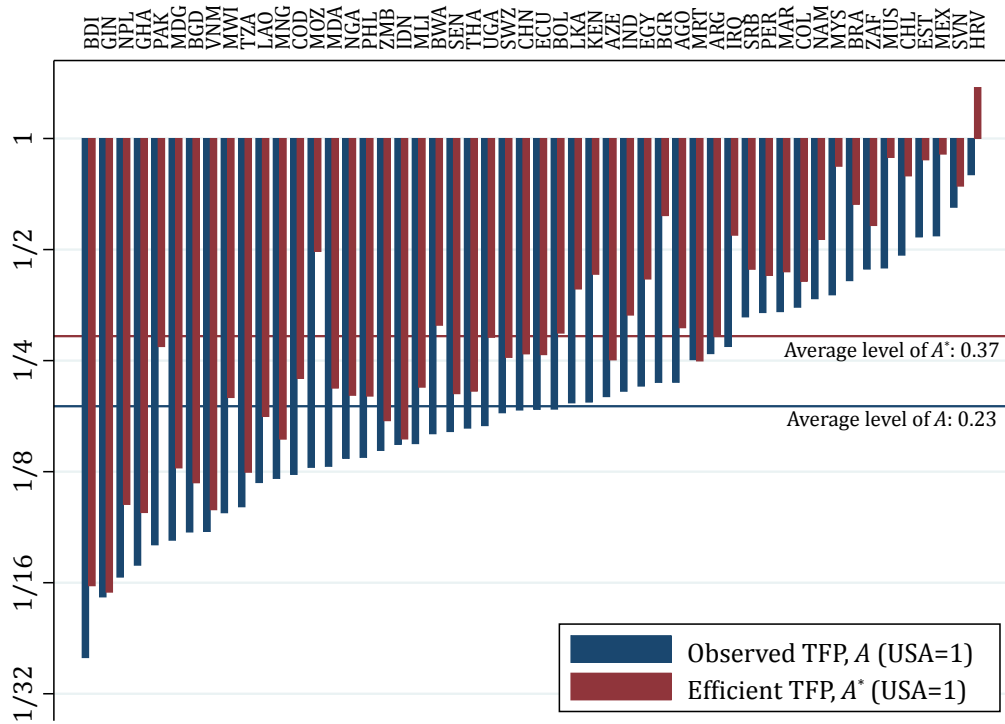
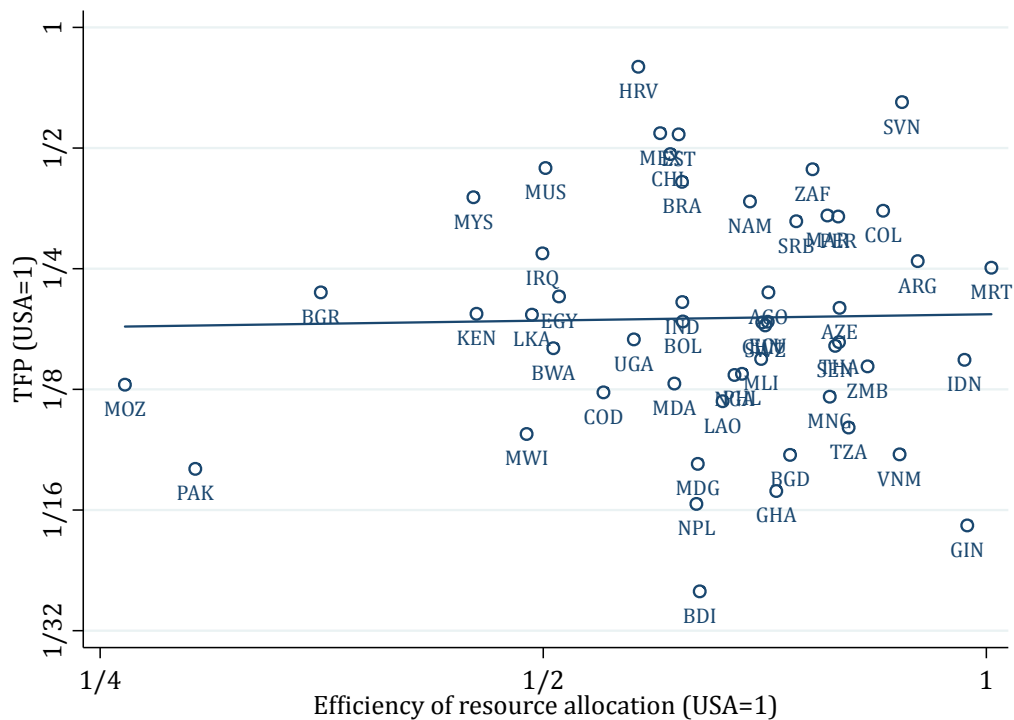


Figure 2 already hints at the second result, namely that the potential gains from removing distortions are not clearly related to the observed TFP level. This is illustrated more specifically in Figure 3, which plots observed TFP against the efficiency of resource allocation. The figure shows that there is no systematic relationship between the efficiency of resource allocation and observed TFP levels (correlation: 0.02). In other words, the poorest countries do not gain most from improving the efficiency of resource allocation. The corollary of this finding (cf. equation (15)) is that efficient TFP levels are highly correlated with observed TFP levels (0.92). We next show that these two results are robust to alternative measures of manufacturing productivity and the efficiency of resource allocation.



## V. SENSITIVITY ANALYSIS

Our first result (cf. Figure 2) is that efficient TFP levels are much higher than observed TFP levels, but that substantial TFP differences remain. To assess the sensitivity of this result, we consider three alternative approaches to determining efficient TFP levels by changing some of the assumptions in the Hsieh and Klenow (2009) model. In the first alternative, we allow output elasticity parameters to vary across countries and industries instead of using benchmark technology parameters from the US. For each country-industry pair, the average labor share across firms is used as the output elasticity of labor and the output elasticity of capital is calculated as one minus the share of labor. In the second alternative, we consider the more extreme case whereby firms within the same industry could adopt different production structures, for example due to lack of access to technologies, and thus end up with different optimal levels of capital intensity. In this case, our measure of misallocation could overestimate the actual level of misallocation since technological differences are wrongly treated as capital distortions. By allowing the optimal capital/labor ratio to be firm-specific, efficient TFP only differs from observed TFP due to output distortions. The third alternative is to change the elasticity of substitution ( $\sigma$ ) from three to five.<sup>19</sup> Assuming a higher elasticity implies higher output distortions (cf. equation (19)), implying that liberalization leads to a larger reallocation of inputs and a larger TFP gain from such reallocation.

<sup>19</sup> This is still well within the range of empirical estimates, see Broda and Weinstein (2006) and Feenstra and Romalis (2014). On the other hand, those estimates are based on exported and imported products and it is conceivable that the average elasticity of substitution is smaller when also including many products that are not (intensively) traded.

**Table 1, Observed and alternative efficient TFP levels (USA=1)**

	Mean	25 <sup>th</sup> percentile	75 <sup>th</sup> percentile
Observed TFP	0.23	0.13	0.33
Efficient TFP			
Baseline	0.37	0.20	0.47
Country elasticities	0.43	0.22	0.54
Firm K/L ratio	0.31	0.16	0.39
$\sigma=5$	0.52	0.26	0.68

Note: Average and percentiles are computed across 52 countries. See the main text for an explanation of the alternatives.

Table 1 presents the results for the various alternative efficient TFP levels and shows that the baseline increase in TFP levels, from an average of 23 to 37 percent, is fairly conservative. Relying on country-specific elasticities leads to larger gains, both at the mean and the 25<sup>th</sup> and 75<sup>th</sup> percentiles. This is because country-specific capital shares are on average higher, giving greater weight to capital distortions. Allowing for firm-specific optimal capital/labor ratios leads to lower gains, since only output distortions remain. Finally, a higher elasticity of substitution leads to much larger gains, as indicated above. For each of these alternatives, though, large differences in efficient TFP levels remain. Taking the most extreme alternative, assuming  $\sigma=5$ , more than doubles average TFP levels – from 23 to 52 percent of the US level. However, even here the 75<sup>th</sup> percentile is at two-thirds of the US level, which means that large productivity differences remain for a substantial majority of countries.

For our second result, namely the lack of a systematic relationship between (log) observed TFP and (log) efficiency of resource allocation, we consider the four misallocation measures from Table 1, as well as a range of alternative observed productivity level estimates. This was less relevant in the first sensitivity analysis as the average observed TFP level varies between 17 and 26 percent and the efficient TFP levels would thus vary in a similar range as that observed in Table 1. We consider five alternative series of observed TPF levels:

1. Rather than relying on a mix of sources on relative wages, we use the observed media wage from WBES for all 52 countries.<sup>20</sup>
2. Rather than using relative rental prices (equation (16)) to estimate factor inputs per worker, we use relative investment prices (equation (18)).
3. Rather than incorporating relative prices of exported and imported goods, we only use relative prices of consumption and investment (as in e.g. Herrendorf and Valentinyi, 2012) for estimating relative value added per worker.
4. Rather than using estimates of industry relative output prices, value added per worker is converted to a common currency using exchange rates (as in e.g. Rodrik, 2013).

<sup>20</sup> Using the mean rather than the median wage does not lead to different results. Using wage data from UNIDO also does not lead to different results.

5. Rather than using the US cost shares of capital and labor, we use the median cost share from the WBES data for each country.<sup>21</sup>

**Table 2, The correlation between observed TFP and the efficiency of resource allocation**

	Baseline allocation	Country elasticities	Firm K/L ratio	$\sigma=5$
Baseline observed TFP	0.02	0.08	0.01	-0.08
Alternative:				
1. Median WBES wages	0.05	0.10	0.12	-0.03
2. Investment prices	-0.03	0.02	-0.10	-0.11
3. Only domestic prices	0.04	0.09	0.03	-0.06
4. Exchange rates	0.06	0.11	0.10	-0.03
5. Country-specific elasticities	0.06	0.10	0.08	-0.04

Note: see the main text for an explanation of the alternatives. No correlation coefficient is significantly different from zero at the 10 percent level or better. Correlations are computed based on the log-transformed TFP and log-transformed efficiency of resource allocation series, as in Figure 3.

Table 2 shows the correlation between observed TFP levels and observed efficiency of resource allocation for each combination of alternative misallocation and observed TFP series. The 0.02 correlation in the top left cell was illustrated above, in Figure 3. All other correlations are similarly small, with no correlation significantly different from zero at even the 10 percent level. This outcome is not surprising as both the alternative TFP levels and the alternative misallocation measures are highly correlated with the baseline series (0.79 and higher).

## VI. CONCLUSIONS

In this paper, we have used firm-level survey data in combination with new estimates of relative productivity levels in manufacturing to analyze the role of resource misallocation in productivity for a set of 52 developing and emerging economies. By applying the Hsieh and Klenow (2009) model of resource misallocation to a broader set of countries and relating these to observed productivity levels, we have provided new evidence on the importance of resource misallocation relative to other factors influencing observed sector productivity. Regarding the scope of this analysis, it is useful to note that the measure of resource misallocation we use here is a broad one, potentially picking up the effects of not only distortions to resource allocation but also of factors such as failed experimentation with new technology by firms. It is also unclear to what extent our findings for manufacturing can be generalized to resource misallocation within other sectors and we are not taking the efficiency of between-sector resource allocation into account

The first result of this paper is that resource misallocation leads to substantially lower productivity levels in manufacturing across a wide range of developing and emerging

<sup>21</sup> Using a common capital share of 33 rather than 40 percent, as is typical in the development accounting literature, does not lead to different results.

economies. If resources were allocated efficiently, the marginal cost of capital and labor would equal the marginal product in all firms in an industry, allowing the more productive firms to grow at the expense of their less-productive counterparts. In this hypothetical efficient setting, the productivity gap relative to the United States would shrink substantially, but at the same time large productivity gaps would remain: the average manufacturing productivity level would increase from 23 to 37 percent of the US level. Resource misallocation across firms, within industries is thus important yet they not the sole factor in explaining low productivity across developing and emerging economies. This suggests a role for slow technology adoption, human capital externalities, misallocation of resources across sectors or any of the other factors that have been associated with productivity in the literature.

This also has an important bearing on our second main finding, that the efficiency of resource allocation and observed productivity levels are essentially uncorrelated in our sample of countries. This means that the least-productive countries are not necessarily the ones with most to gain from more efficient resource allocation. Our indicator of the efficiency of resource allocation thus has a high information content, ranking countries in a way they would not be ranked using more commonly used measures of economic performance. As a consequence, the productivity levels that would prevail if all resource misallocation were eliminated are highly correlated with observed productivity levels. This is a helpful outcome, as it implies that any variable that correlates with observed productivity can safely be assumed to relate to actual productivity since that correlation is not picking up the effect of resource misallocation.

## References

- Adamopoulos, Tasso and Diego Restuccia (2014) "The size distribution of farms and international productivity differences" forthcoming in *American Economic Review*.
- Aghion, Philippe, George-Marios Angeletos, Abhijit Bannerjee and Kalina Manova (2010), "Volatility and growth: Credit constraints and the composition of investment" *Journal of Monetary Economics* 57: 246-265.
- Ashenfelter, Orley (2012), "Comparing Real Wage Rates" *American Economic Review* 102(2): 617-642.
- Bartelsman, Eric, John Haltiwanger and Stefano Scarpetta (2013), "Cross-Country Differences in Productivity: The Role of Allocation and Selection" *American Economic Review* 103(1): 305-334.
- Basu, Susanto and John G. Fernald (2002), "Aggregate productivity and aggregate technology" *European Economic Review* 46(6): 963-991.
- Broda, Christian and David E. Weinstein (2006), "Globalization and the Gains from Trade" *Quarterly Journal of Economics* 121(2): 541-585.
- Buera, Francisco J., Joseph P. Kaboski and Yongseok Shin (2011), "Finance and Development: A Tale of Two Sectors" *American Economic Review* 101(5): 1964-2002.
- Burstein, Ariel and Gita Gopinath (2013), "International Prices and Exchange Rates" forthcoming in Gita Gopinath, Elhanan Helpman and Kenneth Rogoff (eds.) *Handbook of International Economics, Volume IV*, Elsevier.
- Caselli, Francesco (2005), "Accounting for Cross-Country Income Differences" in Philippe Aghion and Steven Durlauf (eds.) *Handbook of Economic Growth*, 679-741 Elsevier.
- Caves, Douglas W., Laurits R. Christensen and W. Erwin Diewert (1982a), "Multilateral comparisons of output, input and productivity using superlative index numbers" *Economic Journal* 92: 73-86.
- Caves, Douglas W., Laurits R. Christensen and W. Erwin Diewert (1982b), "The economic theory of index numbers and the measurement of input, output, and productivity" *Econometrica* 50: 1392-1414.
- Comin, Diego and Bart Hobijn (2010), "An Exploration of Technology Diffusion" *American Economic Review* 100(5): 2031-2059.
- Diewert, W. Erwin (1976), "Exact and superlative index numbers" *Journal of Econometrics* 4: 114-145.
- Elsby, Michael W. L., Bart Hobijn and Aysegül Şahin (2013), "The decline of the U.S. labor share" forthcoming *Brookings Papers on Economic Activity*.
- Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2013), "The Next Generation of the Penn World Table" *NBER Working Paper*, no. 19255.
- Feenstra, Robert C. and John Romalis (2014), "International Prices and Endogenous Quality" forthcoming *Quarterly Journal of Economics* 129(2).
- Fernald, John G. and Brent Neiman (2011), "Growth Accounting with Misallocation: Or, Doing Less with More in Singapore" *American Economic Journal: Macroeconomics* 3(2): 29-74.

- Hall, Robert E. and Dale W. Jorgenson (1967), "Tax policy and investment behavior" *American Economic Review* 57(3): 391-414.
- Herrendorf, Berthold and Ákos Valentinyi (2012), "Which Sectors Make Poor Countries So Unproductive?" *Journal of the European Economic Association* 10(2): 323-341.
- Heston, Alan (1994), "A brief review of some problems in using national accounts data in level of output comparisons and growth studies" *Journal of Development Economics* 44: 29-52.
- Hooper, Peter (1996), "Comparing manufacturing output levels among the major industrial countries", in OECD (ed.) *Industry Productivity – International Comparison and Measurement Issues*, OECD: Paris.
- Hsieh, Chang-Tai and Peter J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4), 1403-1448.
- Hsieh, Chang-Tai and Peter J. Klenow (2010). Development Accounting. *American Economic Journal: Macroeconomics*, 2(1), 207-223.
- Inklaar, Robert and Marcel P. Timmer (2012), "The Relative Price of Services" forthcoming in *Review of Income and Wealth*.
- Inklaar, Robert and Marcel P. Timmer (2013), "Capital, labor and TFP in PWT8.0" downloadable at [www.ggd.net/pwt](http://www.ggd.net/pwt).
- Inklaar, Robert (2010), "The sensitivity of capital services measurement: Measure all assets and the cost of capital" *Review of Income and Wealth* 56(2): 389-412.
- Jones, Charles I. (2011), "Intermediate Goods and Weak Links in the Theory of Economic Development" *American Economic Journal: Macroeconomics* 3(2): 1-28.
- Jones, Charles I. (2013), "Misallocation, Economic Growth, and Input-Output Economics" in Daron Acemoglu, Manuel Arellano, and Eddie Dekel (eds.) *Advances in Economics and Econometrics, Tenth World Congress, Volume II*, Cambridge University Press: Cambridge, UK.
- Jorgenson, Dale W. and Mieko Nishimizu (1978), "U.S. and Japanese economic growth, 1952-1974: an international comparison" *Economic Journal* 88(352): 707-726.
- Karabarbounis, Loukas and Brent Neiman (2014), "The global decline of the labor share" *Quarterly Journal of Economics* 129 (1): 61-103.
- Lequiller, François and Derek Blades (2006), *Understanding National Accounts*, OECD: Paris.
- Melitz, Marc J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity" *Econometrica* 71(6): 1695-1725.
- Muendler, Marc-Andreas (2009), "Converter from SITC to ISIC" downloadable at <http://econ.ucsd.edu/muendler/html/resource.html>.
- Parente, Stephen L. and Edward C. Prescott (1994), "Barriers to Technology Adoption and Development" *Journal of Political Economy* 102(2): 298-321.
- Restuccia, Diego and Richard Rogerson (2008), "Policy Distortions and Aggregate Productivity with Heterogeneous Plants" *Review of Economic Dynamics* 11(4): 707-720.
- Restuccia, Diego and Richard Rogerson (2013) "Misallocation and productivity" *Review of Economic Dynamics* 16(1): 1-10.

- Rodrik, Dani (2013), "Unconditional Convergence in Manufacturing" *Quarterly Journal of Economics* 128(1): 165-204.
- Sørensen, Anders and Bertel Schjerning (2008), "Productivity measurement in manufacturing and the expenditure approach" *Review of International Economics* 16(2): 327-340.
- Syverson, Chad (2008), "Markets: ready-mixed concrete" *Journal of Economic Perspectives* 22(1): 217-234.
- Syverson, Chad (2011), "What determines productivity?" *Journal of Economic Literature* 49(2): 326-365.
- Valentinyi, Ákos and Berthold Herrendorf (2008) "Measuring Factor Income Shares at the Sector Level" *Review of Economic Dynamics* 11(4): 820-835.
- Van Biesebroeck, Johannes (2009), "Disaggregate productivity comparisons: sectoral convergence in OECD countries" *Journal of Productivity Analysis* 32: 63-79.
- Vollrath, Dietrich (2009), "How important are dual economy effects for aggregate productivity?" *Journal of Development Economics* 88(2): 325-334.
- World Bank (2006) *Where is the wealth of nations? Measuring capital for the 21<sup>st</sup> century*, World Bank: Washington D.C.
- World Bank (2008) *Global purchasing power parities and real expenditures – 2005 international comparison program*, World Bank: Washington D.C.



## Appendix

### Estimating industry output prices

The challenge to accurately estimating manufacturing output prices can best be illustrated in a supply and use framework. Suppose that there are  $i=1,\dots,N$  manufactured goods that can be used for final consumption and investment or as intermediate inputs. Furthermore, there is a set of countries  $j=1,\dots,C$ . Our aim is to compare the price of manufacturing output in country  $j$  relative to another country  $k$ . If we would have data on the output value and prices of individual products  $i$ , an estimate of the relative price of manufacturing output would only require aggregating over the relative prices of the individual products, denoted by  $p_{ij}^y$ . The Törnqvist index is such an aggregator function and a flexible one (Diewert, 1976; Caves et al. 1982b):<sup>22</sup>

$$(A1) \quad \log(p_j^y) - \log(p_k^y) = \sum_{i=1}^N \frac{1}{2} (s_{ij}^y + s_{ik}^y) \left( \log(p_{ij}^y) - \log(p_{ik}^y) \right), \text{ where } s_{ij}^y = \frac{p_{ij}^y y_{ij}}{\sum_i p_{ij}^y y_{ij}}$$

Equation (A1) states that the log relative price of manufacturing output is equal to the weighted-average relative output price of individual products, where the weight is the share of each product in overall output, averaged across the two countries under comparison. When the comparison is across more than two countries, the final index would depend on the choice of base country  $k$ . To avoid this, Caves, Christensen and Diewert (1982a) proposed comparing each country not to an actual country but to a (synthetic) average country:

$$(A2) \quad \log(p_j^y) - \overline{\log(p^y)} = \sum_{i=1}^N \frac{1}{2} (s_{ij}^y + \bar{s}_i^y) \left( \log(p_{ij}^y) - \overline{\log(p_i^y)} \right),$$

where the upperbar indicates an arithmetic average across countries. This approach is typically referred as the GEKS method and is used by the OECD, Eurostat and World Bank in their relative price computations.<sup>23</sup>

The problem in implementing equation (A2) is that we do not have reliable data on relative industry output prices for a large sample of countries, and especially not for developing economies. To see how results based solely on the commonly-used expenditure prices are related to the relative output prices, consider the equality between the value of supply and use:<sup>24</sup>

$$(A3) \quad p_{ij}^q q_{ij} + p_{ij}^x x_{ij} + p_{ij}^z z_{ij} = p_{ij}^m m_{ij} + p_{ij}^y y_{ij}, \forall i.$$

<sup>22</sup> A flexible aggregator function is a second-order approximation to an arbitrary twice differentiable linearly homogeneous function.

<sup>23</sup> The only difference is that those organizations would use a Fisher index, rather than a Törnqvist.

<sup>24</sup> Ignoring net taxes on products, which should be added to the right-hand side.

Here  $q$  denotes domestic final demand,  $x$  is exports,  $z$  is intermediate demand,  $y$  is output and  $m$  is imports. As expressed in equation (A3), the value of the supply of each product (shown on the right-hand side) should equal to the value of its demand (on the left-hand side). Next consider prices  $p_{ij}^q, p_{ij}^x, p_{ij}^z, p_{ij}^y$  and  $p_{ij}^m$  for goods  $i$  in each country  $j$ . In this general setting, we allow the price to differ according to each source of supply or use destination.

Next we sum across all products and rearrange:

$$(A4) \quad \sum_{i=1}^N p_{ij}^y y_{ij} = \sum_{i=1}^N (p_{ij}^q q_{ij} + p_{ij}^z z_{ij} + p_{ij}^x x_{ij} - p_{ij}^m m_{ij})$$

The left-hand side shows the total value of manufacturing output, which consists of the sum of domestic final demand, domestic intermediate demand and exports, and subtracts imports. Equation (A4) implies that the relative price of manufacturing output can be either measured directly, as in equation (A2), or indirectly using prices of domestic final and intermediate demand and export and import prices. The indirect alternative to the direct approach of equation (A2) can be expressed as:

$$(A5) \quad \log(p_j^y) - \overline{\log(p^y)} = \sum_{i=1}^N \frac{1}{2} (s_{ij}^y + \bar{s}_i^y) \left[ \tilde{v}_{ij}^q (\log(p_{ij}^q) - \overline{\log(p_i^q)}) + \tilde{v}_{ij}^z (\log(p_{ij}^z) - \overline{\log(p_i^z)}) \right. \\ \left. + \tilde{v}_{ij}^x (\log(p_{ij}^x) - \overline{\log(p_i^x)}) - \tilde{v}_{ij}^m (\log(p_{ij}^m) - \overline{\log(p_i^m)}) \right],$$

where  $\tilde{v}_{ij}^q = \frac{1}{2} \left( \frac{p_{ij}^q q_{ij}}{p_{ij}^y y_{ij}} + \frac{p_i^q q_i}{p_i^y y_i} \right)$ , the share of domestic final expenditure in the value of output of each product, averaged between country  $j$  and the arithmetic mean of shares across all countries, analogous to the definition of  $\bar{s}_i^y$ . The other  $\tilde{v}$ 's are defined analogously.

Equation (A5) allows us to relate the standard approach, which relies solely on relative prices of domestic final expenditure  $q$  to this more comprehensive approach. The standard approach is only valid if either all relative prices are equal to each other or if the share of domestic final expenditure in total output is equal to one. In the more-common case where the share is less than one, another potential bias is when the share of a product in domestic final expenditure,  $p_{ij}^q q_{ij} / \sum_i p_{ij}^q q_{ij}$ , is used rather than the share in output,  $s_i^y$ .

For a broad group of countries, we have implemented a modified version of equation (A5). Specifically, we use data on relative prices of domestic final expenditure from the 2005 ICP round, which covers 146 countries. We supplement that with data on relative prices of exports and imports from Feenstra and Romalis (2014). Data on prices of

domestic intermediate demand are not separately available and we deal with this in two ways. First, part of domestic intermediate demand is supplied from foreign sources, i.e. imports. As imports of intermediate inputs have no (direct) bearing on output prices of domestic producers, imported intermediates can be excluded from the set of intermediate products and imports in equation (A5). Second, we assume that the price that producers charge to domestic final users is equal to the price charged to domestic intermediate users, so  $p^q = p^z$ . Also, we assume that relative *purchaser* prices of domestic final expenditure are equal to relative *producer* prices. In other words, we assume that the trade and transportation margins are equal across countries. Relaxing this stringent assumption would require detailed input-output tables, which are missing for many of the countries we analyze (see also the discussion below). These assumptions lead to the following modified version of equation (A5):

$$(A6) \quad \log(p_j^y) - \overline{\log(p^y)} = \sum_{i=1}^N \frac{1}{2} (s_{ij}^y + \bar{s}_i^y) \left[ \left( \tilde{v}_{ij}^q + \hat{v}_{ij}^z \right) \left( \log(p_{ij}^q) - \overline{\log(p_i^q)} \right) + \tilde{v}_{ij}^x \left( \log(p_{ij}^x) - \overline{\log(p_i^x)} \right) - \hat{v}_{ij}^m \left( \log(p_{ij}^m) - \overline{\log(p_i^m)} \right) \right]$$

where  $\hat{v}_{ij}^z$  denotes the share of *domestic* intermediate demand in output and  $\hat{v}_{ij}^m$  the imports of products for final demand. By implementing equation (A6), we resolve an issue in the literature that has long been known, but never resolved in a satisfactory manner (see e.g. Hooper, 1996). As discussed in Herrendorf and Valentinyi (2012, 330), the price of final demand,  $p^q$  reflects the price of domestically produced goods and of imports, which are produced using ‘world market’ technology. The fact that certain products are imported rather than domestically produced suggests that domestic technology is at least no better than world market technology. This effect would imply that the variation in  $p^q$  will be lower than the variation in  $p^y$ . At the same time, following the logic of the Melitz (2003) model, only the most-productive firms in an economy will export and their prices would have to be competitive in world markets. Which of these effects dominates is hard to say *ex ante*, so it requires implementing equation (A6).

This implementation requires not only data on relative prices, but also on output, domestic (final and intermediate) demand, exports and imports (of final products) by manufacturing product. For most advanced economies and a growing number of emerging economies, such information is available from input-output tables. However, such data is not available for many of the countries we analyze here. We therefore constructed a dataset by combining industry output data from UNIDO, export and import data from Comtrade and domestic final expenditure data from ICP. This requires detailed matching across different product classifications, dealing with missing data and reconciling conflicting data, all of which is discussed in more detail below.

### **Input-output and employment data construction**

Implementing equation (A6) requires data for the left-hand side and right-hand side of equation (A6). This means we need data on the value of (gross) output for individual products and total manufacturing and information on domestic demand, export demand and imports of final products for the same products. With sufficiently detailed input-output tables for each country analyze, this would be fairly straightforward. However, those are not available for the large majority of countries so we combine and reconcile the data sources that are available. For output data we use the UNIDO INDSTAT databases, for domestic final demand we use the ICP basic heading expenditure data and for exports and imports we use the UN Comtrade database. Note that we have no independent information on domestic intermediate demand, so we compute it as a residual.

#### *Product/industry classification and correspondences*

The product/industry classification that we use distinguishes 14 manufacturing industries that together comprise all of manufacturing and is based on the ISIC revision 3 classification system, see Table A1. This is also the classification used in the World Input-Output Database (WIOD) and represents a compromise between a detailed view of manufacturing and limits to data availability.

**Table A1, Product/industry classification**

Product/Industry	ISIC rev. 3 code
Food, beverages and tobacco	15-16
Textiles, textile products and wearing apparel	17-18
Leather and leather products	19
Wood, paper, printing and publishing	20-22
Petroleum and coal products	23
Chemicals, rubber and plastics	24-25
Non-metallic mineral products	26
Basic and fabricated metal products	27-28
Machinery	29
Electronic and optical equipment	30-33
Transport equipment	34-35
Miscellaneous manufacturing	36-37

Much of the gross output data is already available in the ISIC rev. 3 classification. Where the previous, revision 2, system is used, the official correspondence table is used.<sup>25</sup> The export and import data are collected according to the SITC revision 2 system. The correspondence between SITC and ISIC (both rev. 2) is from Muendler (2009).

The ICP basic heading expenditure data are allocated to manufacturing industries based on category names. So, for example, all food products (rice, fresh milk, sugar, etc.) are allocated to the 'Food, beverages and tobacco' industry. Given that the ICP categories are organized by consumption or investment purpose, this means that the correspondence is not precise. The main problem with precision is the investment category 'metal products and equipment', which includes investment in metal products

<sup>25</sup> For the United Nations classification registry and the official correspondence tables, see: <http://unstats.un.org/unsd/class/>.

(27-28), machinery (29) and electronic and optical equipment (30-33). To avoid a biased allocation, we use the share of imported investment goods to split up this expenditure category.

#### *Gross output data*

As mentioned earlier, we rely on the UNIDO INDSTAT database for data on industry gross output. Where available, we use data for 2005 from the 2012 INDSTAT4 database (based on ISIC rev. 3). However, this covers only 29 of the 52 countries. For a further 16 countries, there is data in either the 2012 INDSTAT4 or the 2006 INDSTAT3 (based on ISIC rev. 2) database but for an earlier or later year. Mostly, the data are for a year in the 2000s, but in a few cases we have to go back further. We use information on value added in total manufacturing from the UN National Accounts Main Aggregates Database to put the data on a comparable 2005 basis.<sup>26</sup>

For the 7 countries that have never been covered in UNIDO, we use the following estimation procedure. For most industries, the output share in total manufacturing does not systematically vary with income level and for these we start off with the median cross-country output share. For 5 industries – food, metal, machinery, electronics and transport equipment – there is such a relationship, with the importance of the food industry declining with (the log of) GDP per capita and the other 4 increasing. For these 5 industries we compute the predicted share given the income level. The shares are then normalized to sum to one. The shares are then multiplied by total manufacturing output, which is based on value added from the UN National Accounts Main Aggregates Database and the median value added to gross output ratio across countries.

#### *Import data*

As discussed in the main text, imports should only cover imports of products for final demand. To make this distinction, we use the Broad Economic Classification (BEC), which groups traded products by final use. This allows us to exclude BEC categories that are typically used as intermediates: materials, parts, etc. We apply the distinction used in the World Input-Output Database (WIOD) and classify BEC categories 111, 121, 21, 22, 31, 322, 42 and 53 as intermediate products and exclude these from the import data (see the UN classification registry for details on the individual codes).

#### *Balancing input-output data*

We have data on gross output from UNIDO, exports and imports from Comtrade and domestic final expenditure from ICP and ideally these would be internally consistent without further adjustments. However, it turns out that often imports would exceed domestic final expenditure or that output is smaller than exports plus domestic final expenditure minus imports. Inconsistencies when mixing sources is not uncommon when compiling National Accounts (see e.g. Heston, 1994, or Lequiller and Blades, 2006) and can be due to measurement error, incorrect product correspondence and differing

---

<sup>26</sup> This assumes that the shares of each industry in manufacturing output is unchanged and that the ratio of manufacturing value added to gross output is unchanged.

concepts. As an example of the latter issue, UNIDO's gross output refers only to the formal manufacturing sector, but domestic final expenditure also covers consumption from informal firms. Similarly, domestic final expenditure is valued at purchaser prices, which includes product taxes, trade and transportation margins; gross output is at basic or producer prices; exports is valued fob (free on board) and imports are cif (cost, insurance, freight). Especially the inclusion of product taxes, trade and transportation margins in domestic final expenditure overestimates the size of domestic final expenditure relative to the other flows. Country-specific input-output tables would (again) be needed to fully resolve this, but in their absence we use information from the US input-output tables. Those tables indicate that expenditure on manufacturing products at producer prices is approximately half of expenditure at purchaser prices, with cross-industry variation between about 40 and 60 percent.

In balancing step 1, we multiply domestic final expenditure by one half. In step 2, we reduce imports to be no larger than domestic final expenditure. Data for 40 countries from the WIOD confirms that this constraint holds when input-output tables are available. This adjustment affects about 37 percent of the country/industry pairs in the countries we analyze. This is a substantial share of observations requiring adjustment, but if we follow the same procedure for WIOD countries, the share of imports in domestic supply shows a correlation of 0.54 with the actual input-output data, compared to a correlation of -0.03 when the adjustment is not made.<sup>27</sup>

In step 3, we ensure that industry gross output covers at least exports plus domestic final expenditure minus imports, i.e. domestic intermediate demand is equal to zero. This adjustment affects 35 percent of the country/industry pairs. We could assume that margins make up less than half of domestic final expenditure, which would lead to a smaller number of observations needing adjustment in step 2. However, that would lead to many more adjustments in step 3, so we struck this balance. More in general, this balancing procedure gives greatest weight to the data on domestic final expenditure as the composition of expenditure across industries is left intact. This implies that any differences between our preferred approach and the standard approach – aggregating domestic final expenditure prices using shares in domestic final expenditure – are not (artificially) driven by the balancing choices we make but instead by the differences in the prices of domestic final expenditure, exports and imports. A further reassuring result is that if WIOD data is used directly, rather than our constructed data, the final manufacturing output price levels never differ by more than 1 percent for the group of countries we consider here.

### *Price aggregation*

Prices of domestic final expenditure, exports and imports are all given at a greater level of detail than the 14 industries we analyze. The same Törnqvist/GEKS procedure

---

<sup>27</sup> Even for the WIOD economies, where data is of arguably higher quality in many cases, measurement error, classification mismatches, etc. lead to imports being larger than domestic final expenditure in more than 20 percent of country/industry pairs using the ICP and Comtrade data.

outlined in equation (A2) is used to aggregate the more detailed prices to the level of the 14 industries. At that point, using the balanced input-output data, equation (A6) can be applied to compute aggregate manufacturing relative price levels.

#### *Employment data*

To estimate the number of manufacturing workers in each country, we draw on a number of sources. For 10 of the 52 countries, the UN National Accounts, Official Country Data provides data on the number of workers (employees and self-employed) in manufacturing. For an additional 20 countries, the ILO publishes employment data. For a further 13 countries, the World Bank's World Development Indicators (WDI) publishes the share of workers in industry, a sector that includes workers in mining, utilities and construction in addition to manufacturing workers. We estimate the share of manufacturing in industry value added using UN National Accounts value data and apply this to estimate the share of manufacturing workers in industry. For the 9 countries where no direct employment data is available we regress the share of manufacturing workers in total employment on the share of manufacturing value added in GDP. We apply the predicted share from this regression to the remaining 9 countries. For all countries, we used total employment from the Penn World Table (PWT) version 8.0 as a control total.

**Table A2, Main Results and Data Description by Country**

	Country	Year	Observations	TFP	Efficiency RA
1	Angola	2006	148	0.22	0.71
2	Argentina	2010	514	0.26	0.90
3	Azerbaijan	2009	62	0.20	0.79
4	Bangladesh	2007	1,199	0.09	0.74
5	Bolivia	2006	162	0.18	0.62
6	Botswana	2006	71	0.16	0.51
7	Brazil	2003	1,360	0.41	0.62
8	Bulgaria	2007	347	0.22	0.35
9	Burundi	2006	71	0.04	0.64
10	Chile	2010	562	0.48	0.61
11	China	2003	1,203	0.18	0.70
12	Colombia	2010	508	0.35	0.85
13	Congo, DR	2006	123	0.12	0.55
14	Croatia	2007	199	0.80	0.58
15	Ecuador	2006	182	0.18	0.71
16	Egypt	2004	538	0.21	0.51
17	Estonia	2009	40	0.54	0.62
18	Ghana	2007	243	0.07	0.72
19	Guinea	2006	78	0.06	0.97
20	India	2002	1,563	0.21	0.62
21	Indonesia	2003	329	0.15	0.97
22	Iraq	2011	405	0.27	0.50
23	Kenya	2007	364	0.19	0.45
24	Lao PDR	2009	99	0.12	0.66
25	Madagascar	2009	88	0.08	0.64
26	Malawi	2005	118	0.10	0.49
27	Malaysia	2002	562	0.38	0.45
28	Mali	2007	232	0.15	0.70
29	Mauritania	2006	57	0.25	1.01
30	Mauritius	2005	86	0.45	0.50
31	Mexico	2010	928	0.54	0.60
32	Moldova	2009	53	0.13	0.61
33	Mongolia	2009	99	0.12	0.78
34	Morocco	2004	691	0.34	0.78
35	Mozambique	2007	240	0.13	0.26
36	Namibia	2006	56	0.37	0.69
37	Nepal	2009	59	0.06	0.64
38	Nigeria	2007	849	0.14	0.67
39	Pakistan	2002	670	0.08	0.29
40	Peru	2010	443	0.34	0.79
41	Philippines	2003	526	0.14	0.68
42	Senegal	2007	194	0.16	0.79
43	Serbia	2009	69	0.33	0.74
44	Slovenia	2009	56	0.65	0.88
45	South Africa	2007	591	0.44	0.76
46	Sri Lanka	2004	298	0.19	0.49
47	Swaziland	2006	42	0.18	0.71
48	Tanzania	2006	207	0.10	0.81
49	Thailand	2004	1,242	0.16	0.79
50	Uganda	2006	232	0.17	0.58
51	Vietnam	2005	1,055	0.09	0.87
52	Zambia	2007	265	0.14	0.83

Notes: 'Year' indicates the year in which the WBES survey that we use was conducted; 'Observations' indicates the number of firm observations used; 'TFP' shows the measured TFP level (USA=1); and column 'Efficiency RA' shows the efficiency of resource allocation (USA=1).

**Table A3, Elasticity Parameters for Measuring Misallocation**

<i>Industry (ISIC Code)</i>	Capital share	Labor share
15	0.489	0.511
17	0.306	0.694
18+19	0.234	0.766



20	0.180	0.820
21	0.422	0.578
22	0.346	0.654
24	0.569	0.431
25	0.381	0.619
26	0.311	0.689
27	0.365	0.635
28	0.290	0.710
29	0.281	0.719
31	0.338	0.662
32	0.305	0.695
34+35	0.234	0.766
36	0.291	0.709

Source: Bureau of Labor Statistics (BLS), averaged over the years 2002-2010.

Note: For a few countries in the WBES dataset, the 2-digit industrial classification of industries with ISIC codes 21-22, 25-26 and 27-29 is not known beyond that level of aggregation. In these cases, the average value of the share of labor and capital for the respective 2-digit ISIC industries is used.

**Table A4, Comparison of misallocation calculations with results from Hsieh and Klenow (2009)**

	Hsieh and Klenow (2009)	Current paper
China	86.6	103.0
India	127.5	130.1

Notes: The data used for analysis in the four calculations refer to different years. The calculations of this paper are based on the WBES dataset from the year 2003 for China and 2002 for India. The calculations by Hsieh and Klenow (2009) reported here are based on 2005 data for China and 1994 data for India, which are the latest years covered in their dataset.

